

### **REMARKS**

The Applicant has now had an opportunity to carefully consider the comments set forth in the Office Action that was mailed March 17, 2008. The decision of the reviewers to reopen prosecution and the allowance of **claims 7 and 8** is noted with appreciation. Nevertheless, the rejections of the remainder of the claims according to new grounds are respectfully traversed.

Reexamination and reconsideration of the Application are respectfully requested.

#### **The Office Action**

In the Office Action that was mailed March 17, 2008:

**claims 7 and 8** were allowed;

**claims 2, 4, 5, 10-13, 18 and 22-28** were rejected under 35 USC §103(a) as being unpatentable over U.S. Patent No. 5,963,205 to Sotomayor ("Sotomayor") in view of U.S. Patent Application Publication No. 2001/0042083 A1 by Saito et al. ("Saito");

**claim 6** was rejected under 35 USC §103(a) as being unpatentable over Sotomayor in view of Saito and further in view of U.S. Patent No. 5,276,616 to Kuga et al. ("Kuga");

**claims 15 and 16** were rejected under 35 USC §103(a) as being unpatentable over Sotomayor in view of Saito and further in view of U.S. Patent No. 4,903,229 to Schmidt et al. ("Schmidt");

**claim 17** was rejected under 35 USC §103(a) as being unpatentable over a combination of **four** documents including Sotomayor, Saito, Schmidt and U.S. Patent No. 6,064,397 to Herregods et al. ("Herregods"); and

**claims 20 and 21** were rejected under 35 USC §103(a) as being unpatentable over Sotomayor in view of U.S. Patent No. 5,754,637 to Brooks et al. ("Brooks").

#### **The Present Application**

Briefly, the present application is directed to systems and methods for automatic and semi-automatic document indexing of scanned documents. The invention is useful where a large document is scanned to generate an electronic version of the document. For example, a review of the document may indicate that chapter headings in the

document are rendered in an 18-point font size at a location that is centered horizontally on a page and is two inches from the top of the page. In that case, **a first subsection delimiter** may be defined as any text located two inches from the top of a page and rendered in 18-point font size. Subheadings in the exemplary document might occur anywhere on a page but are rendered in a 16-point font size with underlined characters. Therefore, **a second subsection delimiter** for the document might be defined as underlined 16-point text.

Once one or more subsection delimiters are defined (e.g., by a document processor user), the electronic version of the document is searched to find occurrences of text corresponding to the subsection delimiters. Information regarding each occurrence is used to create an index or table of contents for the document.

### **The Cited Documents**

In stark contrast, the primary reference of the Office Action to Sotomayor is unrelated to providing an index for a document that has to be scanned in order to create an electronic version. Instead, Sotomayor discusses aspects of a word processor and the **much simpler problem** of automatically identifying key topics and phrases in the text of a document **that is being created in the word processor program**. Sotomayor discusses the use of tokens. However, the tokens of Sotomayor are to allow a web browser to use these tokens to format the displayed text for the particular display device of a particular viewer (column 6, lines 5-9). Even if Sotomayor discusses a system of automatic index creation that uses these tokens for the purpose of index creation, Sotomayor does not disclose or suggest a user designating identifying, indicating, selecting, these tokens for use as a delimiter or generating, building or designating a delimiter definition that includes such tokens. The system of Sotomayor is silent with regard to the user or document processing system operator defining a delimiter. The use of tokens in the system of Sotomayor is beyond the control of the user.

It is respectfully submitted that the secondary documents do not cure the deficiencies of Sotomayor.

For example, Saito allegedly discusses a user-defined search template for extracting information from documents (Title). However, it is respectfully submitted that

Saito does not disclose or suggest determining a sub-section delimiter comprises a user indicating a font size, or a user indicating a font style, or a user indicating a text location description, or a user indicating a specific point coordinate within the document.

Instead, Saito discloses a process by which sample documents are processed and characteristics of the documents are recognized and the user is provided an opportunity to name sets of recognized characteristics. That is, the user provides a "user-defined element name" (see Fig. 13). Further in this regard, in the paragraph cited by the Office Action (50), Saito indicates that "the user-defined information is limited to a name" in the illustrated example. Saito also indicates that in an alternative embodiment, other information is provided by a user (lines 10-12). However, in the only other illustrated embodiment, the "other information" is simply a "number of lines" (see Fig. 14) (paragraph 49, lines 19-34). Accordingly, it is respectfully submitted that Saito does not disclose the user designation of font size and font type, etc., for which it is relied and the assertions of the Office Action in this regard represent **clear errors**.

The newly cited reference to Brooks allegedly discloses a document processing system including an imaging unit, a character recognition unit, a dual path module and an encoder. According to Brooks, a document received by the system is sequentially processed through the imaging unit, the character recognition unit, the dual path module and the encoder. The imaging unit images the front face of the document and attempts to identify character data appearing on the face of said document, such as a handwritten courtesy amount appearing on a bank check. The Office Action relies on Brooks for disclosure of performing recognition functions on scanned document data and for alleged disclosure of marking a paper version of the document with at least one predetermined machine-readable demarcation symbol prior to scanning the document. However, Brooks is related to check processing and one concerned with automatic index generation for long scanned documents would not look to Brooks.

Gillings allegedly discusses an electronic document management system that converts documents into electronic images which can be sequentially routed to individual users in a network system. The network system includes **at least two work nodes** for processing the documents where **one of the nodes is a data entry work node**. At the **data entry work node**, data contained in the document is entered into a

database. As the data is entered into the database, it is dynamically linked with its corresponding image (Abstract). Gillings includes the word --indexing-- (e.g., column 10, line 20 - column 11, line 29). However, the indexing process of Gillings involves the manual keying of index information by a worker at a entry node. Accordingly, the combination of Sotomayor with the manual transcription system of Gillings does not arrive at the method for automatic index generation of the present application. It is submitted that the workers of Gillings MANUALLY create the index.

Kuga allegedly discloses a system for creating an index of textual data. In the system, a dictionary stores sets of specialized words particular to a field of knowledge related to the textual data. An entry selecting module selects as index entries only those strings which match one of those specialized words and notes the locations of each occurrence of each index entry in the text. A printer outputs the selected index entries together with their occurrence positions (Abstract).

The Office Action relies on Kuga for disclosure of generating an index for a document with found items corresponding to subsection delimiter occurrences. However, it is respectfully submitted that the dictionary or list of specialized words particular to a field of knowledge of Kuga is not a delimiter definition as the phrase is used in the present application. As used in the present application a delimiter definition is a description of aspects or characteristics of text, not specific text strings. While it is true that text can be used as an aspect of a delimiter definition (for example, the word --chapter-- can be an aspect or characteristic included in the delimiter definition), and that an occurrence corresponding to the delimiter definition might include such text (for example, --Chapter IV: New Hope-- might be an occurrence of a delimiter definition including the word --chapter--) the desired text string itself (e.g., Chapter IV: New Hope) is not a delimiter definition as the phrase is used in the present application. Furthermore, having to generate a list of every word that is desired to be included in an index or table of contents is just the sort of task the subject matter of the present application is directed at avoiding (e.g., page 1, lines 26-30).

Schmidt discloses a forms generating and information retrieval apparatus comprising a compact disc for storing machine readably a plurality of form files and magnetic media for storing machine readably a plurality of information files. Schmidt is

unconcerned with generating an index for a document.

Herregods discloses a method for creating multiple documents having identical background regions and page-specific information regions. Herregods is unconcerned with generating an index or a table of contents for a document.

### **The Claims Are Not Obvious**

**Claims 2, 4, 5, 10-13, 18 and 22-28** were rejected under 35 USC §103(a) as being unpatentable over Sotomayer in view of Saito.

However, with regard to **claims 2, 4, 5 and 22-27**, the Office Action stipulates that Sotomayer fails to disclose determining a sub-section delimiter as recited, for example, in **claim 2** of the present application and relies on Saito, paragraph 50, lines 3-6, for this disclosure.

However, as indicated above, paragraph 50 of Saito clearly indicates that “the user-defined information is **limited to a name**.” While paragraph 50 also asserts that in an alternative embodiment, other information is provided by a user, Saito does not disclose or suggest determining a sub-section delimiter comprises a user indicating at least one of a font size, a font style, a text string, a text location description, and a specific point coordinate within the document or wherein determining a sub-section delimiter comprises a user placing a predetermined machine-readable symbol representing a demarcation point on a printed version of the document as the sub-section delimiter, as is recited, for example, in **claim 2**. The coordinates, indentation, and font type mentioned in paragraph 50 are determined by the system of Saito and are not indicated by a user.

For at least the foregoing reasons, Sotomayer and Saito do not include the subject matter for which they are relied, and **claim 2**, as well as **claims 4 and 5**, which depend therefrom, is not anticipated and is not obvious in light of Sotomayer and Saito.

**Claim 22** recites *inter alia*: wherein determining the sub-section delimiter comprises a user indicating at least one of a font size and a font style. As indicated above, while paragraph 50 mentions a font size and a font type, paragraph 50 **does not disclose or suggest** that a user indicates the font size or the font type. The Office Action stipulates that Sotomayer does not disclose or suggest this subject matter.

Accordingly, the rejection of **claim 22** represents a **clear error** of the Office Action and **claim 22** is not anticipated and is not obvious in light of Sotomayor and Saito.

**Claim 22** recites subject matter similar to that recited in **claim 23** but further recites *inter alia* a user indicating a specific point coordinate within the document. It is respectfully submitted that the Office Action does not assert that Sotomayor and Saito disclose or suggest a user indicating a specific point coordinate within the document.

For at least the foregoing reasons, the rejection of **claim 23** represents a **clear error** of the Office Action and **claim 23** is not anticipated and is not obvious in light of Sotomayor and Saito.

**Claims 24** and **25** depend from **claims 22** and **23**, respectively, and are not anticipated and are not obvious for at least those reasons.

**Claims 26** and **27** depend from **claim 10** and are not anticipated and are not obvious for at least those reasons. Additionally, **claim 26** recites the delimiter designator is operative to accept an indication of at least one of a font size and a font style as a delimiter designation. As indicated above, the Office Action relies on Saito for disclosure of a font size and a font style being used as a delimiter designation, and Saito does not disclose or suggest a delimiter designator module operative to communicate with a document processor operator through a user interface in order to **receive a designation** of a font size or a font style as a delimiter designation, as is recited in the combination of **claims 10** and **26**.

For at least the foregoing additional reasons, **claim 26** is not anticipated and is not obvious in light of Sotomayor and Saito.

**Claim 27** recites subject matter similar to that recited in **claim 26** but further includes reference to a predetermined machine-readable symbol and a specific point coordinate within the document as a delimiter designation. It is respectfully submitted that the Office Action does not assert that Sotomayor and Saito disclose or suggest a delimiter designator operative to accept an indication of a predetermined machine-readable symbol and a specific point coordinate within the document from a document processor operator as a delimiter designation. Accordingly, the rejection of **claim 27** represents a **clear error** of the Office Action, and **claim 27** is not anticipated and is not obvious in light of Sotomayor and Saito.

Furthermore, **the Office has not met its burden** of presenting a *prima facie* case of obviousness. The Office Action implies that one of ordinary skill in the art would have been motivated at the time of the invention to integrate the disclosure of Saito into the disclosure of Sotomayor in order to create a table or index including all the titles of the document as well as the title sections. However, it is respectfully submitted that Sotomayor accomplishes this goal on its own (column 6, lines 7-9, lines 43-49; Fig. 10, presentation page 150; column 10, lines 3-6; column 15, lines 61-65). Furthermore, Sotomayor discloses that the six levels of heading information bracketed by six different heading tokens are sufficient for generating the index of Sotomayor (Abstract) and does not disclose or suggest any need or desire for any other kind of key topic designator. Accordingly, there is no motivation to combine subject matter from Saito into the systems and methods of Sotomayor.

It is respectfully submitted that the only motivation to make such a combination is information gleaned only from the present application. Accordingly, the rejection of **claims 2, 4, and 5** are based on impermissible hindsight reasoning and represent **clear errors** of the Office Action.

For at least the foregoing additional reasons, **claims 2, 4, 5 and 22-27** are not anticipated and are not obvious in light of Sotomayor and Saito.

With regard to **claims 10-13**, the Office Action stipulates that Sotomayor fails to disclose four of the six elements recited in **claim 10**. In this regard, the Office Action asserts that it would have been obvious to an artisan of ordinary skill in the art at the time the invention was made to integrate the disclosure of Saito into the disclosure of Sotomayor such that any paper document can be transformed into an electronic version and stored into a database for easy retrieval as well as avoid losing documents.

However, the documents of Sotomayor are created in electronic form, and Sotomayor does not need assistance from Saito to store electronic versions of documents. Sotomayor is aware of the issues of storing documents in databases and does not require assistance from Saito in this regard (e.g., column 7, lines 11-12). Additionally, it is respectfully submitted that the suggestion to combine Saito and Sotomayor such that any paper document can be transformed into an electronic version and easily retrieved is based only on information gleaned from the present application.

Accordingly, the rejection is based on **impermissible hindsight reasoning**. Saito stores documents for easy retrieval on its own and does not require any assistance from Sotomayor in this regard. Accordingly, there is no motivation to integrate Saito and Sotomayor other than information gleaned from the present application.

For at least the foregoing reasons, **the Office has not met its burden** of presenting a *prima facie* case of obviousness and **claims 10-13** are not anticipated and are not obvious in light of Sotomayor and Saito.

Additionally, **claim 11** recites *inter alia*: the delimiter designator is operative to accept an indication of at least one of a font size, a font style, a text string, a text location description, a predefined machine-readable symbol and a specific point coordinate within the document as a delimiter designation. It is respectfully submitted that Saito does not disclose or suggest a delimiter designator operative to accept this information as a delimiter designation. Instead, the user of the system of Saito is allowed to enter only a user-defined element name and, in some embodiments, a number of lines (Fig. 13; Fig. 14; paragraph 49, lines 20-34; paragraph 50, lines 10-12).

It is respectfully submitted that the user-defined element name is not a delimiter designation. Even if the number of lines is considered to be a delimiter designation, the number of lines is not fairly construed as a font size, a font style, a text string, a text location description, a predefined machine-readable symbol or a specific point coordinate within the document as are recited in **claim 11**.

For at least the foregoing additional reason, **claim 11** is not anticipated and is not obvious in light of Sotomayor and Saito.

With regard to **claims 18 and 28**, the Office Action stipulates that Sotomayor fails to disclose *inter alia* defining a sub-section delimiter as recited in **claim 18**.

In this regard, the Office Action appears to rely on paragraph 50, lines 3-6, of Saito. However, the cited portion while indicating that a search template includes a user-defined element name, their corresponding coordinates, indentation, font size, font type and a number of lines, **does not disclose or suggest** that a document processor operator enters or selects any of this information other than the user-defined element name.

Further in this regard, it is respectfully submitted that the user-defined element



name of Saito is not a component of a delimiter definition. Instead, it is merely a tag or label identifying a definition. Paragraph 50, lines 10-12, indicates that the user-defined information in the disclosed example is limited to a name. **Accordingly, Saito stipulates that the coordinates, indentation, font size and font type mentioned earlier in the paragraph are not entered or defined by the document processor operator.** Therefore, Sotomayor and Saito do not disclose or suggest defining a sub-section delimiter, wherein defining the sub-section delimiter includes a document processor building a sub-section delimiter definition from a list of predetermined potential sub-section delimiter components. Sotomayor and Saito **do not disclose or suggest** entering a sub-section delimiter by selecting symbols on a displayed portion of the electronic version of the document. Even if Saito discloses selecting areas which are highlighted by the processing system of Saito, this **does not disclose or suggest** selecting symbols on a displayed portion of an electronic version of a document. Sotomayor and Saito **do not disclose or suggest** designating at least one demarcation point on a least one displayed portion of the electronic document to create a list of demarcation points to be used as a set of delimiter definitions. Even if the search template of Saito includes location information, the templates of Saito include more than that information. Accordingly, Saito does not disclose or suggest a document processor operator designating at least one demarcation point on at least one displayed portion of the electronic document to create a list of demarcation points to be used as a set of delimiter definitions.

Additionally, Sotomayor and Saito **do not disclose or suggest** a document processor operator entering a sub-section delimiter through keyboard keystrokes. It is respectfully submitted that the user-defined element name of Saito is not a sub-section delimiter. The user-defined element name is simply a tag or label applied to a template. Additionally, it is respectfully submitted that the number of lines is not a sub-section delimiter entered through keyboard keystrokes. It is respectfully submitted that it is unclear what purpose the number of lines parameter is envisioned as serving in the system of Saito. For that matter, it is respectfully submitted that Saito does not explain what any of the parameters are used for.

For example, according to the present application, a font size and a font style are

combined as two tests to determine whether a portion of a document of interest meets the delimiter definition criteria. Saito does not explain whether the font size and font type of, for example, Fig. 15 are to be combined in this way. Moreover, Saito does not explain what the purpose of the number of lines parameter is. In Fig. 14, the selected area is three lines long, yet the number of lines indicated at reference numeral 710 is only two. Saito does not explain how this number is to be used. Moreover, Saito does not disclose or suggest that the number of lines is to be used as a sub-section delimiter entered through keyboard keystrokes as recited in **claim 18**.

For at least the foregoing reasons, **claim 18** is not anticipated and is not obvious in light of Sotomayor and Saito.

The Office Action does not address the recitation of defining a sub-section delimiter of **claim 28**. Accordingly, the rejections of **claims 18** and **28** represent an inappropriate omnibus rejection and a clear error of the Office Action.

To the extent that **claim 28** recites subject matter similar to that recited in **claim 18**, arguments similar to those submitted in support of **claim 18** are submitted in support of **claim 28**. Additionally, it is respectfully submitted that Sotomayor and Saito **do not disclose or suggest** performing statistical analysis on recognized characters to select characteristics that are most likely to be associated with sub-section delimiters.

For at least the foregoing reasons, **claim 28** is not anticipated and is not obvious in light of Sotomayor and Saito.

Additionally, even if Saito is to be construed as disclosing some aspect of the sub-section delimiter definition elements of **claims 18** and/or **28**, it is respectfully submitted that there is **no motivation in the art** to combine the subject matter from Saito into subject matter from Sotomayor other than information gleaned only from the present application. Accordingly, the rejection of **claims 18** and **28** are based on impermissible hindsight reasoning.

In the explanation of the rejection of **claim 10**, the Office Action asserts that the motivation for combining subject matter from Sotomayor with subject matter of Saito is for easy retrieval and to avoid losing documents. However, Sotomayor is aware of and addresses the storage of documents in databases and does not disclose or suggest a need for assistance from Saito in this regard. Accordingly, the Office has not met its

burden of presenting a *prima facie* case of obviousness, and **claims 18** and **28** are not anticipated and are not obvious in light of Sotomayor and Saito.

**Claim 6** was rejected in light of three documents including Sotomayor, Saito and Kuga.

However, **claim 6** depends from **claim 2** and is not anticipated and not obvious for at least that reason.

Additionally, there is **no motivation in the art** for combining Sotomayor, Saito and Kuga. The lack of motivation for combining Sotomayor and Saito has been addressed above. Additionally, Sotomayor and Saito do not disclose or suggest any chances for erroneous results. Accordingly, there is no motivation in the art to combine any disclosure of editing text that might be included in Kuga with the subject matter of Sotomayor and Saito.

The only motivation for doing so is information gleaned only from the present application. Accordingly, the rejection of **claim 6** is based on impermissible hindsight reasoning.

For at least the foregoing additional reasons, **claim 6** is not anticipated and is not obvious in light of Sotomayor, Saito and Kuga.

**Claims 15** and **16** were rejected under 35 USC §103(a) in light of three documents including Sotomayor, Saito and Schmidt. However, **claims 15** and **16** depend from **claim 10** and are not anticipated and are not obvious for at least that reason.

**Claim 17** was rejected under 35 USC §103(a) as being unpatentable over a combination of four documents including Sotomayor, Saito, Schmidt and Herregods. However, **claim 17** depends from **claim 15**, which depends in turn from **claim 10**. Accordingly, **claim 17** is not anticipated and is not obvious for at least those reasons.

**Claims 20** and **21** were rejected under 35 USC §103(a) as being unpatentable over Sotomayor in view of Brooks.

However, **Brooks is related to the art of check processing and one looking to solve problems related to automatic and semi-automatic index generation for large scanned documents** would not look to Brooks. Accordingly, **Brooks is nonanalogous art**, and there is no motivation to combine Sotomayor and Brooks.

Additionally, Brooks discusses using character recognition to ascertain a courtesy amount written on a check and is not fairly interpreted as disclosing defining a sub-section delimiter by marking a paper version of a document with at least one predetermined machine-readable demarcation symbol prior to scanning the document as recited in **claim 20**.

Furthermore, it is respectfully submitted that the Office has not met its burden for presenting a *prima facie* case of obviousness. There is **no motivation in the art** to combine the aspects of the check processing system of Brooks into the word processor system of Sotomayor.

It is respectfully submitted that the assertion that it would be obvious to integrate disclosure of Brooks into the disclosure of Sotomayor such that the producer or writer of the document can designate the sub-section delimiter before the document is transformed into an electronic version such that the sub-section delimiter is not changed by an unauthorized person, and therefore, ensures the authenticity of the document and reserves the rights for the writers is specious.

Brooks discloses a system for processing checks after they have been filled in order to make a withdrawal or draft from a checking account. It is respectfully submitted **the only motivation to modify the system of Brooks to incorporate a document authoring tool**, as appears to be the assertion or the suggestion of the Office Action, is **information gleaned only from the present application**.

Alternatively, there is no disclosure or suggestion to combine any aspect of a check processing system into the word processor with an automatic index creation aspect of Sotomayor. It is respectfully submitted that the only motivation to do so is information gleaned from the present application. Accordingly, the rejection of **claims 20** and **21** is based on impermissible hindsight reasoning, and **claims 20** and **21** are not anticipated and are not obvious in light of Sotomayor and Brooks.

Furthermore, **claim 21** depends from **claim 2** and is not anticipated and is not obvious for at least that reason.

Additionally, the rejection of **claims 20** and **21** does not appear to address the subject matter of **claim 21**. Accordingly, the rejections of **claims 20** and **21** appear to be an inappropriate, omnibus-type rejection and therefore represents a **clear error** of

the Office Action.

Moreover, it is respectfully submitted that Sotomayor and Brooks do not disclose or suggest the automatically generated index is an automatically generated table of contents of the document, and the items corresponding to the defined sub-section delimiter are chapter titles displayed in an order in which they appear in the document. In this regard, it is respectfully submitted that each check or document of Brooks includes only one courtesy amount. Accordingly, Brooks does not disclose or suggest the automatic generation of a table of contents wherein the items corresponding to the defined sub-section delimiter are chapter titles as recited in **claim 21**.

For at least the foregoing additional reasons, **claim 21** is not anticipated and is not obvious in light of Sotomayor and Brooks.

#### **Telephone Interview**

In the interests of advancing this application to issue the Applicant(s) respectfully request that the Examiner telephone the undersigned to discuss the foregoing or any suggestions that the Examiner may have to place the case in condition for allowance.

## CONCLUSION

**Claims 2, 4-8, 10-13, 15-18 and 20-28** remain in the application. For at least the foregoing reasons, the application is in condition for allowance. Accordingly, an early indication thereof is respectfully requested.

☒ Remaining Claims, as delineated below:

(1) FOR	(2) CLAIMS REMAINING AFTER AMENDMENT LESS HIGHEST NUMBER PREVIOUSLY PAID FOR	(3) NUMBER EXTRA
TOTAL CLAIMS	23	- 29 =
INDEPENDENT CLAIMS	8	- 8 =

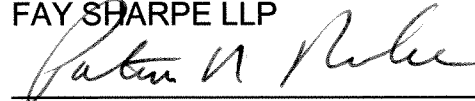
☒ This is an authorization under 37 CFR 1.136(a)(3) to treat any concurrent or future reply, requiring a petition for extension of time, as incorporating a petition for the appropriate extension of time.

☒ The Commissioner is hereby authorized to charge any filing or prosecution fees which may be required, under 37 CFR 1.16, 1.17, and 1.21 (but not 1.18), or to credit any overpayment, to Deposit Account 24-0037.

In the event the Examiner considers personal contact advantageous to the disposition of this case, he/she is hereby authorized to call Thomas Tillander, at Telephone Number (216) 861-5582.

Respectfully submitted,

FAY SHARPE LLP



Patrick R. Roche, Reg. No. 29,580  
Thomas Tillander, Reg. No. 47,334  
1100 Superior Avenue, Seventh Floor  
Cleveland, OH 44114-2579  
216-861-5582

June 4, 2008  
Date